A STATISTICAL ANALYSIS OF LEAST-SQUARES CIRCLE-CENTRE ESTIMATION

Emanuel E. Zelniker and I. Vaughan L. Clarkson

Intelligent Real-Time Imaging and Sensing Group
School of Information Technology & Electrical Engineering
The University of Queensland
Queensland, 4072, AUSTRALIA
{zelniker,v.clarkson}@itee.uq.edu.au

ABSTRACT

In this paper, we examine the problem of fitting a circle to a set of noisy measurements of points on the circle's circumference. An estimator based on standard least-squares techniques has been proposed by DELOGNE which has been shown by KÅSA to be convenient for its ease of analysis and computation. Using CHAN's circular functional model to describe the distribution of points, we perform a statistical analysis of the estimate of the circle's centre, assuming independent, identically distributed Gaussian measurement errors. We examine the existence of the mean and variance of the estimator for fixed sample sizes. We find that the mean exists when the number of sample points is greater than 2 and the variance exists when this number is greater than 3. We also derive approximations for the mean and variance for fixed sample sizes when the noise variance is small. We find that the bias approaches zero as the noise variance diminishes and that the variance approaches the CRAMÉR-RAO lower bound. We also show this through Monte-Carlo simulations.

1. INTRODUCTION

The accurate fitting of a circle to noisy measurements of points on its circumference is an important and much-studied problem in statistics. It has applications in many areas of research including archæology [1], geodesy [2], microwave engineering [3] and computer vision and metrology [4].

The problem of obtaining an accurate circular fit, by which we mean the estimation of a circle's centre and its radius, appears to have been first studied by THOM [1] in connection with measurements of ancient stone circles in Britain. He proposes an approximate method of least squares solution. In addressing a problem of 'statistical geography', ROBINSON [2] gives a complete formulation of the solution to the problem by the method of least squares.

The first detailed statistical analysis to be published appears to be that of CHAN [5]. He proposes a 'circular functional relationship', which we also use as the basis for our investigations. In this model, it is assumed that the measurement errors are instances of independent and identically distributed (i.i.d.) random variables. Additionally, the points are assumed to lie at fixed but unknown angles around the circumference, *i.e.*, not only are the centre and radius of the circle unknown parameters to be estimated, but so are the angles of each circumferential point. He derives a method to find the maximum likelihood estimator (MLE) when the errors have a Gaussian distribution. This method is identical to the least-squares method of [2]. He also examines the consistency of the estimator.

A disadvantage of the MLE is that it is difficult to analyse. From a numerical point of view, another disadvantage is that the only known algorithms for computing the MLE are iterative. Furthermore, it has been reported that there are instances in which there is no minimum, but rather a stationary point, or several local minima in the likelihood function [6]. The difficulties with the MLE were recognised by KÅSA [3], who proposes using a simple estimator due to DELOGNE [7] which is relatively easy to analyse and also to compute and is based on least-squares techniques. This estimator has subsequently been independently rediscovered at least twice [8, 9].

BERMAN & CULPIN [6] have carried out a statistical analysis of the MLE and the DELOGNE-KÅSA estimator (DKE). Specifically, they prove some results regarding the asymptotic consistency and variance of the estimates. CHAN & THOMAS [10] have investigated the CRAMÉR-RAO lower bound (CRLB) for estimation in the circular functional model, but see also [11].

In this paper, we are interested in the properties of the DKE for fixed (small) sample sizes rather than its asymptotic properties. KÅSA himself carries out a 'first-order' or 'small-error' analysis of the estimator [3]. However, when the random variables that give rise to the errors are Gaussian, it can no longer be guaranteed that the errors will always be small, no matter how small the variance, and so the analysis becomes invalid. At the outset, it is not even clear whether the mean or variance of the estimator exists.

KÅSA states, by way of justifying the first-order analysis, that 'it may be appreciated that [the expressions for the mean and variance of the estimator] are, in general, very hard to evaluate'. Nevertheless, in this paper, we demonstrate that, under certain conditions, the defining integrals are not wholly intractable. From analysis of the integrals, we set out conditions for which the mean and variance of the DKE for circle centre exist for fixed sample sizes under CHAN's circular functional model with Gaussian errors. Where the mean exists, we show that the estimator is unbiased in the limit as noise variance approaches zero. Where the variance exists, we show that the variance approaches the CRLB as the variance approaches zero. We additionally demonstrate these theoretical results through Monte-Carlo simulations.

Emanuel Zelniker is additionally supported by a scholarship from the Commonwealth Scientific & Industrial Research Organisation.

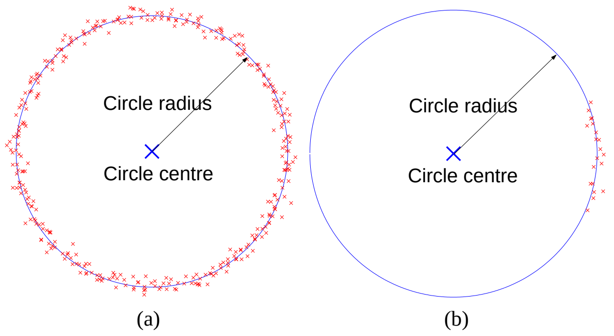


Fig. 1. Two examples of noisy measurements of points on the circumference of a circular arc.

2. BACKGROUND

2.1. Chan's Circular Functional Model

We now briefly present CHAN's circular functional model [5]. In this model, we assume that the positions of N points on the circumference of a circle are measured. The measurement process introduces random errors so that the Cartesian coordinates (x_i, y_i), $i = 1, \dots, N$ can be expressed as

$$x_i = a + r\cos\theta_i + \xi_i, \quad y_i = b + r\sin\theta_i + \eta_i.$$

Here, (a, b) is the centre of the circle, r is its radius, the θ_i are the angles around the circumference on which the points lie and the ξ_i and η_i are instances of random variables representing the measurement error. They are assumed to be zero-mean and i.i.d. In addition, we will specify that they are Gaussian with variance σ^2. In this paper, we explicitly exclude the possibility that $r = 0$ or $\theta_1 = \theta_2 = \ldots = \theta_N$.

Figure 1 shows some data with N points for the circumference of a circular arc, $(x_1, y_1), \dots, (x_N, y_N)$, displaced from the circumference by noise.

2.2. Maximum Likelihood Estimation

If we define $r_i(a, b) = \sqrt{(x_i - a)^2 + (y_i - b)^2}$, then CHAN [5] showed that the MLE is

$$(\hat{a}_{\text{ML}}, \hat{b}_{\text{ML}}, \hat{r}_{\text{ML}}) = \arg\min_{(a,b,r)} \sum_{i=1}^{N} [r_i(a, b) - r]^2.$$

The difficulties with the MLE are that it is hard to analyse and also to compute numerically. Analytically, it is not certain that a global minimum exists, or whether there might be local minima [6]. Numerically, the only methods available for solution are iterative. This raises the usual issues with convergence and sensitivity to the initial solution estimate.

2.3. The Delogne-Kåsa Estimator

The analytical and numerical difficulties with the MLE in Section 2.2 led KÅSA [3] to propose the use of a modified estimator, originally due to DELOGNE [7] which we can write as

$$(\hat{a}_{\text{DK}}, \hat{b}_{\text{DK}}, \hat{r}_{\text{DK}}) = \arg\min_{(a,b,r)} \sum_{i=1}^{N} [r_i^2(a, b) - r^2]^2.$$

The linearisation which results from this formulation simplifies the analysis and the computation considerably. It can be shown that this estimator is a standard linear least-squares estimator. In terms of matrix algebra, we have

$$\mathbf{Z} = (\hat{a}_{\text{DK}}, \hat{b}_{\text{DK}})^T = \tfrac{1}{2}(\mathbf{S} + \mathbf{T})^{\#}(\mathbf{u} + \mathbf{v}). \tag{1}$$

Here, the superscript '#' represents the MOORE-PENROSE generalised inverse or pseudo-inverse and, for a matrix $\mathbf{A}$ where $\mathbf{A}^T\mathbf{A}$ is non-singular, we may write $\mathbf{A}^{\#} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$. Further,

$$\begin{aligned} \mathbf{S} &= (\mathbf{s}_1, \mathbf{s}_2) = \mathbf{P}\mathbf{S}', \ \mathbf{T} = \mathbf{P}\mathbf{T}', \\ \mathbf{u} &= \mathbf{P}\mathbf{u}' \ \text{ and } \ \mathbf{v} = \mathbf{P}\mathbf{v}' \end{aligned} \tag{2}$$

where

$$\mathbf{S}' = \begin{pmatrix} a + r\cos\theta_1 & b + r\sin\theta_1 \\ \vdots & \vdots \\ a + r\cos\theta_N & b + r\sin\theta_N \end{pmatrix}, \mathbf{T}' = \begin{pmatrix} \xi_1 & \eta_1 \\ \vdots & \vdots \\ \xi_N & \eta_N \end{pmatrix}.$$

The elements of $\mathbf{u}'$ and $\mathbf{v}'$ are given by the expressions

$$\begin{aligned} u_i' &= (a + r\cos\theta_i)^2 + (b + r\sin\theta_i)^2, \\ v_i' &= 2(a + r\cos\theta_i)\xi_i + 2(b + r\sin\theta_i)\eta_i + \xi_i^2 + \eta_i^2 \end{aligned}$$

and $\mathbf{P}$ is an $N \times N$ symmetric projection matrix defined so that $\mathbf{P} = \mathbf{I} - (\mathbf{1}\mathbf{1}^T/N)$ where $\mathbf{I}$ is the $N \times N$ identity matrix and $\mathbf{1}$ is an N-dimensional column vector, all of whose entries are 1. Note that $\|\mathbf{P}\|_2 = 1$.

3. ANALYSIS OF THE DELOGNE-KÅSA ESTIMATOR

We now turn our attention to the analysis of the DKE for fixed sample sizes. We are firstly interested in the question of whether the mean and variance exist. Later, we derive low-variance approximations for their values which are valid whenever they exist.

Before outlining the proofs for the main theorems in this section, we observe the following lemmas. The full proofs of all theorems and lemmas may be found in [12].

Lemma 1. *The matrix $\mathbf{P}$ has a singular-value decomposition of the form*

$$\mathbf{P} = \mathbf{\Upsilon}\mathbf{\Delta}\mathbf{\Upsilon}^T$$

where $\mathbf{\Upsilon}$ is an orthogonal matrix and $\mathbf{\Delta} = \operatorname{diag}\{1, \dots, 1, 0\}$.

Lemma 2. *For any vectors $\mathbf{x}, \boldsymbol{\mu} \in \mathbb{R}^N$,*

$$\exp\left(-\frac{\|\mathbf{x} - \boldsymbol{\mu}\|_2^2}{2\sigma^2}\right) \le \exp\left(\frac{\|\boldsymbol{\mu}\|_2^2}{2\sigma^2}\right) \exp\left(-\frac{\|\mathbf{x}\|_2^2}{4\sigma^2}\right).$$

Corollary 1. *If $\mathbf{X} = (X_1, \dots, X_N)^T$ is a multivariate normal random vector such that each $X_i \sim N(\mu_i, \sigma^2)$ is independent, then*

$$E[\|f(\mathbf{X})\|_2^k] \le 2^{N/2} \exp\left(\frac{\|\boldsymbol{\mu}\|_2^2}{2\sigma^2}\right) E[\|f(\mathbf{Y})\|_2^k],$$

where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)^T$ and $\mathbf{Y} = (Y_1, \dots, Y_N)^T$ is a multivariate normal random vector such that each $Y_i \sim N(0, 2\sigma^2)$ is independent.

Definition 1. *We say that an* $N \times n$ *matrix* $\mathbf{X}$ *is a rectangular Gaussian matrix if each element is i.i.d. with identical variance* σ^2 *and* $E[\mathbf{X}] = \boldsymbol{\mu}$. *We denote its distribution* $G(N, n, \boldsymbol{\mu}, \sigma^2)$.

Theorem 1. *The mean of the DKE for circle centre, as defined in (1), exists if the number of sample points on the circumference,* N, *is greater than* 2.

Outline of proof. If the variance σ^2 is zero then $\mathbf{Z}$ is deterministic and $\mathbf{Z} = \frac{1}{2}\mathbf{S}^{\#}\mathbf{u}$. In this case, the mean clearly exists, since the pseudo-inverse of $\mathbf{S}$ always exists. Hence, we restrict our attention to the case where $\sigma^2 > 0$.

In order to show that the expectation exists, it is necessary to show that $E[\|\mathbf{Z}\|_2] < \infty$. Let $\mathbf{Y}'$ be a random matrix with distribution $G(N, 2, \mathbf{S}', \sigma^2)$ such that $\mathbf{Y}' = \mathbf{S}' + \mathbf{T}'$. From the definition of expectation,

$$\begin{aligned} E[\|\mathbf{Z}\|_2] &= \frac{1}{2} E[\|(\mathbf{S}+\mathbf{T})^{\#}(\mathbf{u}+\mathbf{v})\|_2] \\ &= \frac{1}{2} E[\|(\mathbf{P}\mathbf{Y}')^{\#}[\mathbf{u}+\mathbf{v}(\mathbf{Y}')]\|_2] \end{aligned} \tag{3}$$

and we note that $\mathbf{v}$ is a function of $\mathbf{Y}'$, *i.e.*, a function of the ξ_i and η_i. Through the use of Corollary 1 and the sub-multiplicative inequality, we find that

$$\begin{aligned} & E[\|\mathbf{Z}\|_2] \\ & \le 2^{\frac{N-2}{2}} \exp\left(\frac{\|\mathbf{S}'\|_F^2}{2\sigma^2}\right) E[\|(\mathbf{P}\mathbf{W}')^{\#}\|_2 \|\mathbf{u}+\mathbf{v}(\mathbf{W}')\|_2], \end{aligned} \tag{4}$$

where $\|\cdot\|_F$ represents the FROBENIUS norm of its argument and $\mathbf{W}'$ is a random matrix like $\mathbf{Y}'$ but each element has zero mean and twice the variance, *i.e.*, $\mathbf{W}' \sim G(N, 2, \mathbf{0}, 2\sigma^2)$.

Define $\boldsymbol{\delta}$ and $\boldsymbol{\epsilon}$ as the first $N-1$ elements of $\boldsymbol{\Upsilon}^T\boldsymbol{\xi}$ and $\boldsymbol{\Upsilon}^T\boldsymbol{\eta}$, as defined in Lemma 1. Let us also define $\bar{\xi}$ and $\bar{\eta}$ to be the last (*i.e.*, N^{th}) elements of $\boldsymbol{\Upsilon}^T\boldsymbol{\xi}$ and $\boldsymbol{\Upsilon}^T\boldsymbol{\eta}$ so that

$$\begin{pmatrix} \boldsymbol{\delta} \\ \bar{\xi} \end{pmatrix} = \boldsymbol{\Upsilon}^T\boldsymbol{\xi}, \quad \begin{pmatrix} \boldsymbol{\epsilon} \\ \bar{\eta} \end{pmatrix} = \boldsymbol{\Upsilon}^T\boldsymbol{\eta}. \tag{5}$$

Notice that $\boldsymbol{\Upsilon}^T\mathbf{W}' \sim G(N, 2, \mathbf{0}, 2\sigma^2)$. Then, by Lemma 1

$$\boldsymbol{\Delta}\boldsymbol{\Upsilon}^T\mathbf{W}' = \begin{pmatrix} \mathbf{W} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{W} = (\boldsymbol{\delta}, \boldsymbol{\eta}),$$

and $\mathbf{W} \sim G(N-1, 2, \mathbf{0}, 2\sigma^2)$. Also,

$$\|(\mathbf{P}\mathbf{W}')^{\#}\|_2 = \|\mathbf{W}^{\#}\|_2. \tag{6}$$

We would now like to note that it is not very difficult to bound $\|\mathbf{u}+\mathbf{v}(\mathbf{W}')\|_2$ above by a degree two polynomial in $\|\mathbf{W}\|_2$, $|\bar{\xi}|$ and $|\bar{\eta}|$. Since these are independent random variables, we can take the expectation of $|\bar{\xi}|$ and $|\bar{\eta}|$ separately so that we can now rewrite (4) to show that

$$E[\|\mathbf{Z}\|_2] \le 2^{\frac{N-2}{2}} \exp\left(\frac{\|\mathbf{S}'\|_F^2}{2\sigma^2}\right) E[\|\mathbf{W}^{\#}\|_2 p_1(\|\mathbf{W}\|_2)], \tag{7}$$

where $p_1(\|\mathbf{W}\|_2)$ is a degree two polynomial in $\|\mathbf{W}\|_2$ only.

Consider the value of $\|\mathbf{W}\|_2$ and $\|\mathbf{W}^{\#}\|_2$, *i.e.*,

$$\begin{aligned} \|\mathbf{W}\|_2 &= s_{\text{max}}, \\ \|\mathbf{W}^{\#}\|_2 &= \frac{1}{s_{\text{min}}}. \end{aligned} \tag{8}$$

Now, s_{max} and s_{min} are the singular values of $\mathbf{W}$ and therefore are the square roots of the maximum and minimum eigenvalues of $\mathbf{W}^T\mathbf{W}$. Then, $\mathbf{W}^T\mathbf{W}$ has a WISHART distribution and from MUIRHEAD [13, p. 106], the exact joint density function for the 2 singular values of $\mathbf{W}^T\mathbf{W}$ can be written as

$$\begin{aligned} & P_{s_{\text{max}}, s_{\text{min}}}(s_{\text{max}}, s_{\text{min}}) \\ & = \begin{cases} K_{N,2} \exp\left(-\frac{(s_{\text{max}}^2 + s_{\text{min}}^2)}{2}\right) s_{\text{max}}^{N-2} s_{\text{min}}^{N-2} (s_{\text{max}}^2 - s_{\text{min}}^2) & \\ & \text{if } s_{\text{max}} \ge s_{\text{min}}, \\ 0 & \text{otherwise}, \end{cases} \end{aligned} \tag{9}$$

where $K_{N,2}$ is a normalising constant. Looking at (7) and using (8), we can say that

$$\begin{aligned} E[\|\mathbf{W}^{\#}\|_2 p_1(\|\mathbf{W}\|_2)] = \frac{1}{(2\pi\sigma^2)} \int_0^\infty \int_0^\infty \frac{p_1(s_{\text{max}})}{s_{\text{min}}} \\ P_{s_{\text{max}}, s_{\text{min}}}(s_{\text{max}}, s_{\text{min}})\, ds_{\text{max}}\, ds_{\text{min}}, \end{aligned} \tag{10}$$

and by substituting (9) into (10), we have that

$$\begin{aligned} E[\|\mathbf{W}^{\#}\|_2 p_1(\|\mathbf{W}\|_2)] = \frac{K_{N,2}}{(2\pi\sigma^2)} \int_0^\infty \int_0^\infty p_1(s_{\text{max}}) \\ \left(\frac{1}{s_{\text{min}}^2} - \frac{1}{s_{\text{max}}^2}\right) s_{\text{max}}^N s_{\text{min}}^{N-1} \\ \exp\left(-\frac{(s_{\text{max}}^2 + s_{\text{min}}^2)}{2}\right) ds_{\text{max}}\, ds_{\text{min}}. \end{aligned} \tag{11}$$

It can now be seen from (11) that we have bounded $E[\|\mathbf{Z}\|_2]$ above a two-dimensional integral in s_{max} and s_{min}. This integral is the product of a degree 2 polynomial of non-negative powers of s_{max} and s_{min} with an exponential of the negative square of s_{max} and s_{min} when $N \ge 3$. Such an integral is finite, *e.g.*, see [14, §3.461]. $\square$

For the remainder of this section, we omit the proofs of our theoretical results. Instead, in the next section, we will demonstrate them through simulation results. The full proofs of all theorems and lemmas may be found in [12], and all follow a similar theme to the proof of Theorem 1.

Theorem 2. *The variance of the DKE for the circle centre, as defined in (1), exists if the number of sample points on the circumference,* N, *is greater than* 3.

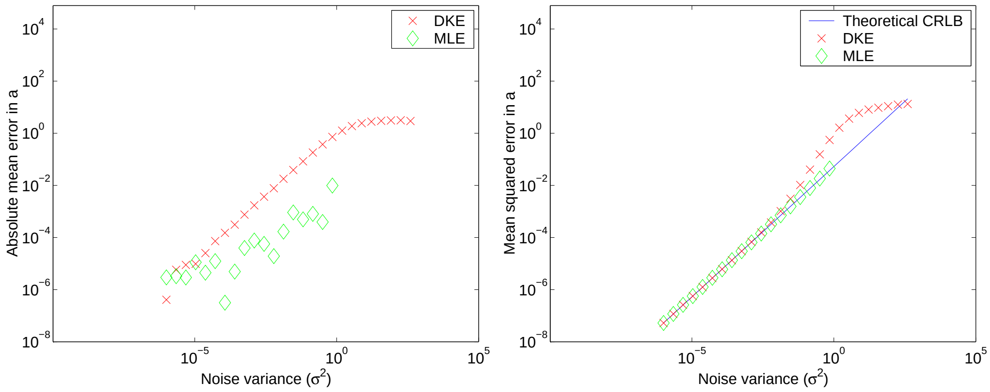
Theorem 3. *When the mean of the DKE exists,*

$$E[\mathbf{Z}] = (a, b)^T + O(\sigma).$$

Theorem 4. *When the variance of the DKE exists,*

$$\text{var}(\mathbf{Z}, \mathbf{Z}) = r^2\sigma^2(\mathbf{S}^T\mathbf{S})^{-1} + O(\sigma^3). \tag{12}$$

We note that it is not difficult to show that the expression $r^2\sigma^2(\mathbf{S}^T\mathbf{S})^{-1}$ in (12) is equal to the upper 2-by-2 sub-matrix of $\mathbf{J}^{-1}$, where $\mathbf{J}$ is the FISHER Information Matrix defined in [10, 11]. This is the CRLB for CHAN's model. Hence, the DKE for circle centre approaches statistical efficiency for fixed sample sizes as the noise approaches zero.



(a) (b)

Fig. 2. Simulation results for the DKE and MLE for varying σ for an arc length of π radians.

4. RESULTS FROM SIMULATION

The DKE was simulated using a Monte-Carlo analysis. In each trial, 200 points ($N = 200$) were generated in equal increments around half a circle's circumference. The radius r was set to 5. Then, noise was added to each (x_i, y_i) coordinate pair in the form of (ξ_i, η_i). The amount of noise, σ was varied from 10^{-3} to 20 in equal geometric increments. Then, the DKE was run repeatedly, 15 000 times, for each value of σ to obtain estimates for the centre of the circle $(\hat{a}, \hat{b})$ and use them to generate mean error values and mean square error (MSE) values. The mean error values in $\hat{a}$ are plotted versus σ^2 in Figure 2a on a logarithmic scale. It can be seen that the mean error decreases with decreasing σ. This is consistent with Theorems 1 and 3. The MSE values in $\hat{a}$ are plotted against their corresponding theoretical CRLB for the same level of noise σ in Figure 2b on a logarithmic scale. It can be seen that at high values of noise, the estimator $\hat{a}$ departs from the CRLB. However, as the noise level, σ approaches zero, the estimator $\hat{a}$ approaches the theoretical CRLB. This is consistent with Theorems 2 and 4.

The same procedure was carried out for the MLE (in diamonds) using the NEWTON-RAPHSON algorithm, so it was started from an arbitrary point not too far from the true minimum. It can be seen that for high values of σ, the errors are very large (this is why the last 8 simulation results could not be displayed in both plots).

We have chosen to illustrate the results with plots for $\hat{a}$. Plots for $\hat{b}$ follow an identical pattern.

5. ACKNOWLEDGEMENT

The authors would like to acknowledge the contribution of PROFESSOR BARRY QUINN whose many helpful comments have led directly to the improvement of this paper. We would also like to thank DR. PETER KOOTSOOKOS for providing feedback on drafts of this paper.

6. REFERENCES

[1] A. Thom, "A Statistical Examination of the Megalithic Sites in Britain," *J. Roy. Statist. Soc. Ser. A General*, vol. 118, pp. 275–295, Mar. 1955.

[2] Stephen M. Robinson, "Fitting Spheres by the Method of Least Squares," in *Comm. ACM*, 1961, p. 491.

[3] I. Kåsa, "A circle fitting procedure and its error analysis," *IEEE Trans. Instrum. Meas.*, vol. 25, pp. 8–14, Mar. 1976.

[4] U. M. Landau, "Esimation of a Circular Arc Center and its Radius," *Comput. Vision Graph. Image Process.*, vol. 38, pp. 317–326, June 1986.

[5] Nai Ng Chan, "On Circular Functional Relationships," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 27, pp. 45–56, Apr. 1965.

[6] Mark Berman and David Culpin, "The Statistical Behaviour of Some Least Squares Estimators of the Centre and Radius of a Circle," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 48, pp. 183–196, Oct. 1986.

[7] P. Delogne, "Computer Optimization of Deschamps' method and error cancellation in reflectometry," in *Proc. IMEKO-Symp. Microwave Measurements*, Budapest, Hungary, 1972, pp. 117–123.

[8] Samuel M. Thomas and Y. T. Chan, "A Simple Approach to the Estimation of Circular Arc Center and its Radius," *Comput. Vision Graph. Image Process.*, vol. 45, pp. 362–370, Aug. 1989.

[9] I. D. Coope, "Circle Fitting by Linear and Non-Linear Least Squares," *J. Optim. Theory Appl.*, vol. 76, pp. 381–388, Feb. 1993.

[10] Y. T. Chan and Samuel M. Thomas, "Cramér-Rao Lower Bound for Estimation of a Circular Arc Center and its Radius," in *Graph. Models Image Process.*, Nov. 1995, pp. 527–532.

[11] K. Kanatani, "Cramér-Rao Lower Bounds for Curve Fitting," in *Graph. Models Image Process.*, 1998, pp. 93–99.

[12] E. E. Zelniker and I. V. L. Clarkson, "A Statistical Analysis of the DELOGNE-KÅSA Method for Fitting Circles," Submitted to *IEEE Trans. Signal Process.*, June 2003.

[13] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[14] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals Series and Products*, Academic Press, San Diego, 2000.